Estatística Inferencial

Prof. Wagner Hugo Bonat

Departamento de Estatística
Universidade Federal do Paraná

Principais modelos multivariados

Importância das distribuições multivariadas

- Descrevem o comportamento conjunto de várias variáveis aleatórias.
- Permite analisar de forma simultânea várias variáveis aleatórias.
- Investigar a relação entre elas.
- Usar a informação de uma para inferir sobre a outra.
- Descrever estruturas de dependência temporal, espacial, genética, etc.
- Condensar a informação de várias variáveis em um número reduzido de fatores latentes.

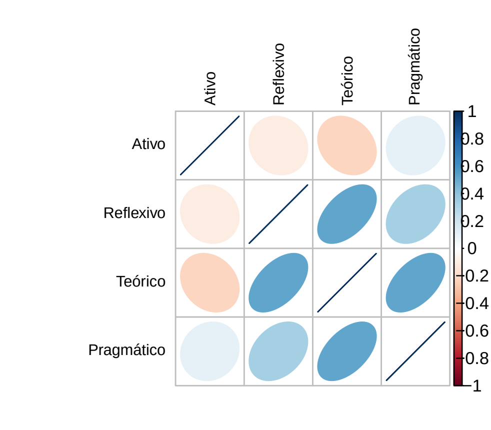


Figura 1. Correlação entre os escores para os estilos de aprendizado determinados para os alunos de Estatística Básica.

A distribuição Multinomial

Características de uma v.a. Multinomial

- **Generalização** da distribuição Binomial.
  - Binomial: 2 resultados mutuamente exclusivos $\{0, 1\}$.
  - Multinomial: k resultados mutuamente exclusivos.
- Condições:
  - Suponha n ensaios **independentes** que podem apresentar, em cada, apenas um de k possíveis resultados $\{O_1, O_2, \ldots, O_k\}$.
  - Sejam $p_1, p_2, \ldots, p_k$, com $\sum p_i = 1$, as probabilidades de observar $O_1, O_2, \ldots, O_k$, respectivamente, que se mantém **constantes**.
  - Seja Y_i o **número de vezes** que o resultado O_i ocorre nos n ensaios.
  - Dessa forma, $\{Y_1, \ldots, Y_k\} \sim \text{Multinomial}(p_1, \ldots, p_k, n)$.

j	O_1	O_2	$\cdots$	O_k
1		✓		
2	✓			
3		✓	$\vdots$	
$\vdots$				✓
n		✓		
Soma	Y_1	Y_2	$\cdots$	Y_k

Exemplos de uma v.a. Multinomial

1. Resultado do arremesso da bola à cesta:
$$\{\text{errou, fez 1, fez 2, fez 3 pontos}\}.$$
2. Ação resultado de uma propaganda direta por email
$$\{\text{não viu, acessou, wishlist, compra}\}.$$
3. Categoria de veículo alugado por cliente em um site
$$\{\text{SUV, Sedan, Hatch, Utilitário}, \ldots\}.$$
4. Classificação de um estudante em relação à área do Curso
$$\{\text{Exatas, Humanas, Biológicas}\}.$$
5. Signo de uma pessoa.
6. Dia da semana de acidente de trabalho.
7. Grau de uma infração de trânsito.
8. Estado civil de uma pessoa.

A distribuição Multinomial

Satisfeitas as condições apresentadas anteriormente, tem-se que a distribuição do vetor aleatório $\{Y_1, \ldots, Y_k\}$ é dada por

$$
\begin{aligned}
p(y_1, \ldots, y_k) &= \frac{n!}{y_1! \cdot \ldots \cdot y_k!} \cdot p_1^{y_1} \cdot \ldots \cdot p_k^{y_k} \\
&= \frac{n!}{\prod_{i=1}^{k} y_i!} \cdot \prod_{i=1}^{k} p_i^{y_i}.
\end{aligned}
$$

j	O_1	O_2	$\cdots$	O_k
1		✓		
2	✓			
3		✓	$\vdots$	
$\vdots$				✓
n		✓		
Soma	Y_1	Y_2	$\cdots$	Y_k

Esperança e variância para cada componente i é dado por

$$
\begin{aligned}
\mathrm{E}(O_i) &= p_i & \text{logo} \quad \mathrm{E}(Y_i) &= n \cdot p_i. \\
\mathrm{V}(O_i) &= p_i(1 - p_i) & \text{logo} \quad \mathrm{V}(Y_i) &= n \cdot p_i(1 - p_i).
\end{aligned}
$$

Exemplo: bolas de gude

Em uma urna há 2 bolas vermelhas, 3 verdes e 5 azuis. São selecionadas 4 bolas ao acaso da urna com reposição. Qual a probabilidade de retirar 2 verdes e 2 azuis.

Figura 2. Bolas de gude. Fonte: https://cutt.ly/QhzaxK8.

Solução

- São $n = 4$ ensaios independentes.
- As probabilidades são:

$$p_1 = \frac{2}{10}, \quad p_2 = \frac{3}{10} \quad \text{e} \quad p_3 = \frac{5}{10}.$$

- Deseja-se a probabilidade do resultado

$$y_1 = 0, \quad y_2 = 2 \quad \text{e} \quad y_3 = 2.$$

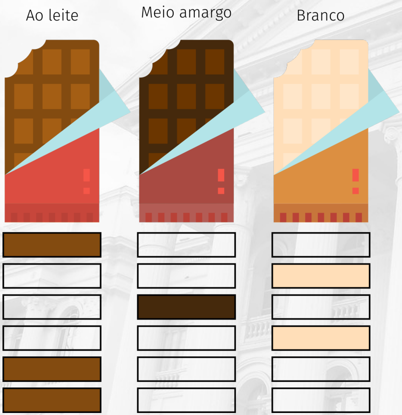
Usando a função de probabilidade, tem-se

$$\begin{aligned} p(0, 2, 2) &= \frac{4!}{0! \cdot 2! \cdot 2!} \cdot 0.2^0 \cdot 0.3^2 \cdot 0.5^2 \\ &= 0.135 \end{aligned}$$

Diferença entre Multinomial e Binomiais

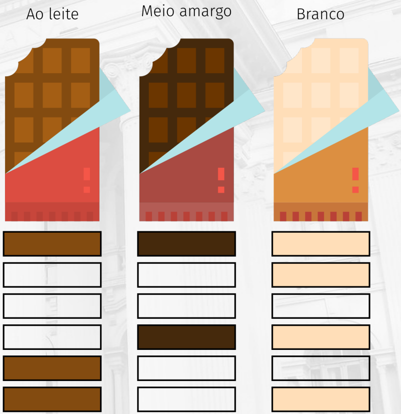
Multinomial

Qual seu chocolate favorito?



Binomiais

Qual chocolate o cliente comprou?



Ícone retirado do https://www.flaticon.com/authors/photo3idea-studio

A distribuição Normal Multivariada

Características de um v.a. Normal Multivariada

- Um conjunto de variáveis contínuas não limitadas.
- Individualmente as variáveis são Normais, portanto simétricas.
- Variáveis apresentam relação linear entre si.
- Resultam da combinação de muitos fatores de pequena contribuição.
  - Características genéticas e morfológicas.
  - Traços latentes.
  - Índices econômicos.

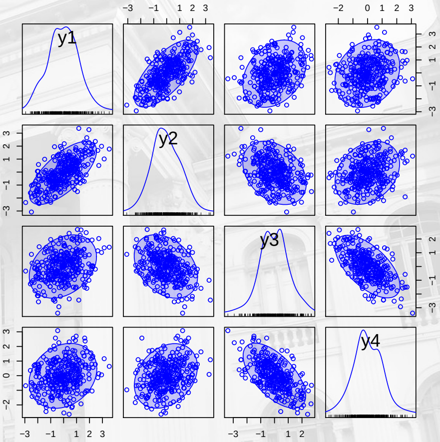


Figura 3. Diagramas de dispersão para 4 variáveis contínuas.

Exemplos de v.a. com distribuição Normal Multivariada

- Peso de massa seca, altura de planta e altura da primeira espiga em uma planta de milho.
- Peso, comprimento e circunferência de uma banana.
- Comprimentos no crânio de um fóssil ou animal.
- Variação de um conjunto de índices econômicos ou mercadorias: e.g. combustível.
- Traço latente para resolução de problemas de física, matemática, química, etc.

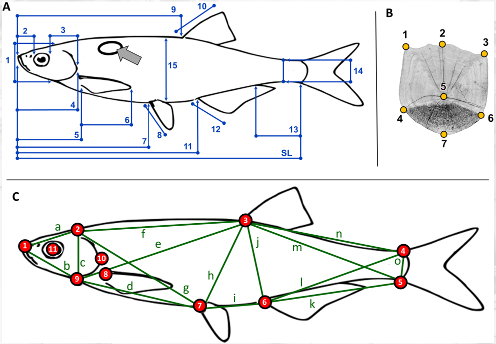


Figura 4. Distâncias medidas em um peixe para estudo morfométrico. Fonte: https://cutt.ly/AhlsOWo.

Importância da Normal Multivariada

- Modelar a estrutura de correlação ou dependência.
  - Modelos de séries temporais: correlação entre datas.
  - Modelos geoestatísticos: correlação entre pontos no espaço.
  - Modelos genéticos: correlação entre características.
- Técnicas de análise multivariada.
  - Análise de discriminante linear.
  - Análise fatorial exploratória e confirmatória.
  - Análise de correlação canônica.
  - Análise de variância multivariada.

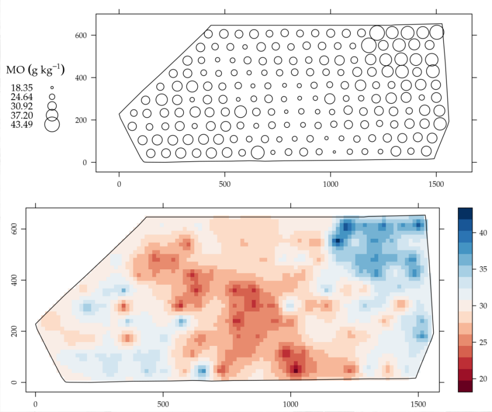


Figura 5. Gráfico de valores observados de matéria orgânica do solo e o mapa com a predição para todo o terreno.

A distribuição Normal Multivariada

Um vetor aleatório contínuo $\mathbf{Y}$ $(p \times 1)$ tem distribuição Normal multivariada se sua densidade conjunta é dada por

$$f(\mathbf{y}) = \left(\frac{1}{2\pi}\right)^{p/2} \det(\Sigma)^{-1/2} \exp\left\{-\frac{1}{2}(\mathbf{y}-\boldsymbol{\mu})^\top \Sigma^{-1}(\mathbf{y}-\boldsymbol{\mu})\right\},$$

em que

$$\mathbf{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_p \end{bmatrix}, \quad \boldsymbol{\mu} = \begin{bmatrix} \mu_1 \\ \vdots \\ \mu_p \end{bmatrix} \quad \text{e} \quad \Sigma = \begin{bmatrix} \sigma_1^2 & \ldots & \sigma_{1p} \\ \vdots & \ddots & \vdots \\ \sigma_{p1} & \ldots & \sigma_p^2 \end{bmatrix}.$$

Denotamos por $\mathbf{Y} \sim \mathrm{N}_p(\boldsymbol{\mu}, \Sigma)$.

A média e variância do vetor aleatório são dados por

$$\mathrm{E}(\mathbf{Y}) = \boldsymbol{\mu} \quad \text{e} \quad \mathrm{V}(\mathbf{Y}) = \Sigma.$$

Distribuições marginais e condicionais (caso bivariado)

Distribuição marginal: A distribuição marginal de cada elemento de um vetor aleatório Normal é também Normal. Isto é,

$$Y_i \sim \mathrm{N}(\mu_i, \sigma_i^2) \quad i \in \{1, 2\}.$$

Distribuição condicional: A distribuição condicional de cada elemento de um vetor aleatório em relação ao outro tambem é Normal. Isto é,

$$Y_i | Y_j = y_j \sim \mathrm{N}(\mu_{i|j}, \sigma_{i|j}^2), \qquad i \neq j \in \{1, 2\},$$

em que

$$\mu_{i|j} = \mu_i + \frac{\sigma_{ij}}{\sigma_j^2}(y_j - \mu_j) \quad \text{e} \quad \sigma_{i|j}^2 = \sigma_i^2 - \frac{\sigma_{ij} \cdot \sigma_{ji}}{\sigma_j^2}.$$

Exemplo: altura de mães e filhas

A altura de mães e filhas adultas segue distribuição Normal bivariada com parâmetros

$$\mu = \begin{bmatrix} 165 \\ 165 \end{bmatrix}, \quad \Sigma = \begin{bmatrix} 49 & 35 \\ 35 & 49 \end{bmatrix}.$$

- Dado que uma mãe tem 172 cm de altura, qual a estatura esperada para a filha quando adulta?
- Qual a probabilidade da filha ter mais de 180 cm?

Figura 6. Mãe com sua filha. Foto de Pixabay no Pexels.

Solução

- A altura esperada para a filha é

$$\mu_{f|m} = \mu_f + \frac{\sigma_{fm}}{\sigma_m^2}(y_m - \mu_m)$$
$$= 165 + \frac{35}{49}(172 - 165) = 170.$$

- A variância de $Y_f|Y_m = 172$ é

$$\sigma^2_{f|m} = \sigma^2_f - \frac{\sigma_{fm} \cdot \sigma_{mf}}{\sigma^2_m}$$
$$= 49 - \frac{35 \cdot 35}{49} = 47.57.$$

- E assim, $P(Y_f > 180|Y_m = 172) = 0.074$.

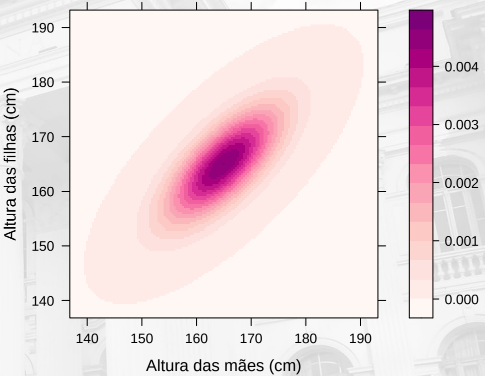


Figura 7. Distribuição Normal bivariada para o problema da altura de mães e filhas.

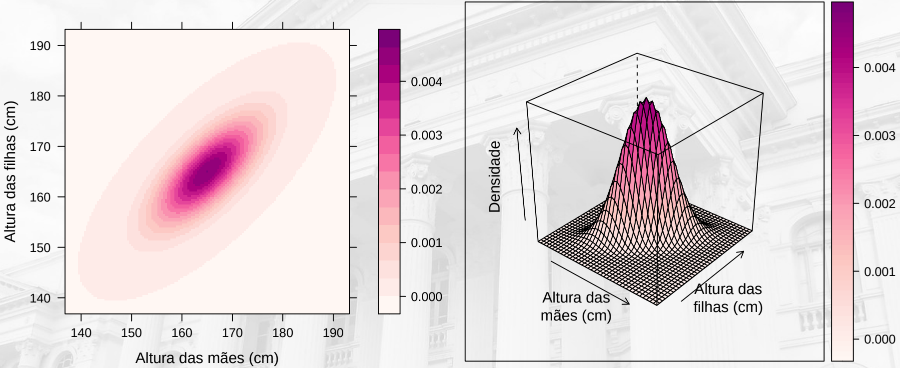


Figura 8. Distribuição Normal bivariada para o problema da altura de mães e filhas.

Distribuições marginais e condicionais (caso geral)

Considere a partição do vetor aleatório $\mathbf{Y}$ em dois subvetores complementares de tamanho p e q. Ou seja,

$$\mathbf{Y} = \begin{bmatrix} \mathbf{Y}_p \\ \mathbf{Y}_q \end{bmatrix}, \quad \boldsymbol{\mu} = \begin{bmatrix} \boldsymbol{\mu}_p \\ \boldsymbol{\mu}_q \end{bmatrix}, \quad \Sigma = \begin{bmatrix} \Sigma_{pp} & \Sigma_{pq} \\ \Sigma_{qp} & \Sigma_{qq} \end{bmatrix}.$$

Distribuição marginal: A distribuição marginal de $\mathbf{Y}_p$ é Normal. Isto é,

$$\mathbf{Y}_p \sim \mathrm{N}_p(\boldsymbol{\mu}_p, \Sigma_{pp}).$$

Distribuição condicional: A distribuição condicional de $\mathbf{Y}_p$ em relação à $\mathbf{Y}_q$ é Normal. Isto é,

$$\mathbf{Y}_p|\mathbf{Y}_q = \mathbf{y}_q \sim \mathrm{N}_p(\boldsymbol{\mu}_{p|q}, \Sigma_{p|q}), \quad \text{em que}$$

$$\boldsymbol{\mu}_{p|q} = \boldsymbol{\mu}_p + \Sigma_{pq}\Sigma_{qq}^{-1}(\mathbf{y}_q - \boldsymbol{\mu}_q) \quad \text{e} \quad \Sigma_{p|q} = \Sigma_{pp} - \Sigma_{pq}\Sigma_{qq}^{-1}\Sigma_{qp}.$$

Algumas propriedades

- Correlação 0 (zero) implica em **independência**.
- Variáveis individualmente apresentarem distribuição Normal não implica em distribuição Normal conjunta.

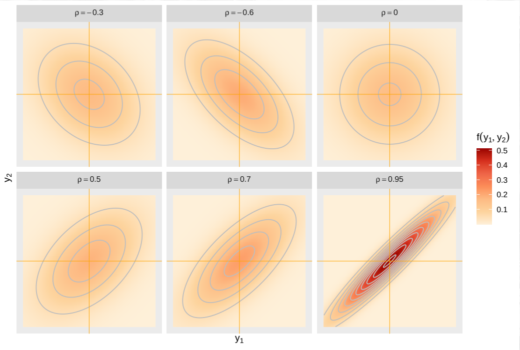


Figura 9. Gráficos da densidade da Normal bivariada com diferentes valores para a correlação.

Teoremas limites

Motivação

- Sejam $Y_1, Y_2, \ldots, Y_n$ variáveis aleatórias (v.a) independentes e idênticamente distribuídas (iid) com $\mathrm{E}(Y_i) = \mu$ e $\mathrm{V}(Y_i) = \sigma^2$.
- Objetivo: Baseado em um conjunto de realizações de Y_i *estimar* o valor μ.
- Estratégia simples: Usar a versão empírica de μ, i.e.

$$\hat{\mu} = \sum_{i=1}^{n} \frac{Y_i}{n}.$$

- $\hat{\mu}$ é uma variável aleatória?
- Qual a esperança de $\hat{\mu}$?
- Qual a variância de $\hat{\mu}$?
- Qual é a distribuição de probabilidade de $\hat{\mu}$?

Sequência de v.a's

- **Definição (Sequência de v.a's):** Sejam $X_1, X_2, \ldots, X_n$ v.a's com função de probabilidade (fp) ou função densidade de probabilidade (fdp) $f(X, \theta)$. Para uma função $g(\cdot)$ define-se uma sequência de v.a's como,

$$\begin{aligned} Y_1 &= g(X_1) \\ Y_2 &= g(X_1, X_2) \\ &\vdots \\ Y_n &= g(X1, \ldots, X_n). \end{aligned}$$

- Estatísticas amostrais são funções do tamanho da amostra e podem ser tratadas como sequência de v.a's.
- Exemplos: Média e variância amostral.

Convergência em probabilidade

- **Convergência em probabilidade**: Seja $Y_1, \ldots, Y_n$ uma sequência de v.a's. Dizemos que Y_n converge em probabilidade para uma constante ou v.a Y, se $\forall \epsilon > 0$, tem-se

$$\lim_{n \to \infty} \mathrm{P}(|Y_n - Y| > \epsilon) = 0.$$

- Notação: $Y_n \xrightarrow{P} Y$.
- Observações:
  1. $Y_n \xrightarrow{P} Y$ é equivalente a $(Y_n - Y) \xrightarrow{P} 0$.
  2. Para vetores aleatórios, tem-se

$$\boldsymbol{Y}_n = (Y_{n1}, \ldots, Y_{nk})^\top, \boldsymbol{Y}_n \xrightarrow{P} \boldsymbol{Y} \quad \text{se}$$

$$Y_{ni} \xrightarrow{P} \boldsymbol{Y}_i \quad \text{para} \quad i = 1, \ldots, n.$$

Desigualdades

Desigualdades

- **Markov:** Seja Y uma v.a não negativa com $\mathrm{E}(Y) = \mu$ e $\epsilon > 0$. Então,

$$\mathrm{P}(Y \geq \epsilon) \leq \frac{\mathrm{E}(Y)}{\epsilon}.$$

- Demonstração.
- Chebyshev: Seja Y qualquer v.a com $\mathrm{E}(Y) = \mu$ e $\mathrm{V}(Y) = \sigma^2$ ambos finitos. Então, para todo $\epsilon > 0$

$$\mathrm{P}(|Y - \mu| > \epsilon) \leq \frac{\mathrm{V}(Y)}{\epsilon^2} = \frac{\sigma^2}{\epsilon^2}.$$

- Demonstração.
- Corolário: Seja $\psi(\cdot)$ uma função monotônica, então

$$\mathrm{P}(\psi(|Y|) > \psi(\epsilon)) \leq \frac{\mathrm{E}(\psi(Y))}{\psi(\epsilon)}.$$

Leis dos grandes números

Lei dos Grandes Números: Chebyshev

- **Teorema**: Seja $Y_1, \ldots, Y_n$ v.a. iid com $\mathrm{E}(Y_i) = \mu$ e $\mathrm{V}(Y_i) = \sigma^2$ ambos finitos. Então,

$$\overline{Y} = \frac{1}{n} \sum_{i=1}^{n} Y_i \xrightarrow{P} \mathrm{E}(Y_i) = \mu.$$

- Demonstração.
- Ilustração computacional.

Lei dos Grandes Números: Kolmogorov

- **Teorema**: Seja $Y_1, \ldots, Y_n$ v.a. iid com $\mathrm{E}(|Y_i|) < \infty$ e $\mathrm{E}(Y_i) = \mu$. Então,

$$\overline{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i \xrightarrow{P} \mathrm{E}(Y_i) = \mu.$$

- Condição $\mathrm{E}(|Y_i|) < \infty$ significa que $\mathrm{E}(|Y_i|) = \int |y_i| f(y_i, \theta) dy_i < \infty$.
- Controla as caudas da distribuição.
- Caudas não podem ser muito pesadas, tal que $\mathrm{E}(|Y_i|) = \infty$, mas ainda podem ser pesadas a ponto de $\mathrm{E}(Y_i^2) = \infty$.
- Kolmogorov não requer que $\mathrm{V}(Y_i)$ exista.
- Kolmororov cobre distribuições com caudas pesadas como a t-Student.
- Demonstração é trabalhosa em sua forma geral.
- Ilustração computacional.

Manipulação de limites e convergência em probabilidade

- **Teorema de Slutsky**: Sejam Y_n e Z_n sequências de v.a's e sejam b, c e d constantes.
  - Se $Y_n \xrightarrow{P} c$ então $bY_n \xrightarrow{P} bc$.
  - Se $Y_n \xrightarrow{P} c$ e $Z_n \xrightarrow{P} d$ então $Y_n + Z_n \xrightarrow{P} c + d$.
  - Se $Y_n \xrightarrow{P} c$ e $Z_n \xrightarrow{P} d$ então $\frac{Y_n}{Z_n} \xrightarrow{P} \frac{c}{d}$ desde que $d \neq 0$ e $Y_n Z_n \xrightarrow{P} cd$.
  - Se $Y_n \xrightarrow{P} c$ e $h(\cdot)$ é uma função contínua então $h(Y_n) \xrightarrow{P} h(c)$.
- Demonstração (ver Magalhães (2006) Ed.2 pag 319).

Preliminares: Teorema Central do Limite

Convergência em distribuição

- Seja $Y_1, \ldots, Y_n$ uma sequência de v.a. Dizemos que Y_n converge em distribuição para uma v.a Y e escrevemos $Y_n \xrightarrow{D} Y$ se

$$\mathrm{F}_{Y_n}(y) = \mathrm{P}(Y_n \leq y) \rightarrow \mathrm{F}_Y(y) = \mathrm{P}(Y \leq y), \quad n \rightarrow \infty.$$

- Na maioria das aplicações Y será normal ou qui-quadrado.
- Convergência em distribuição em geral é demonstrada via Teorema Central do Limite.
- Se n é grande usamos a distribuição de Y como uma aproximação para a distribuição de Y_n.
- Distribuição exata de Y_n em geral é dificil de se obter.

Função geradora de momentos

- **Proposição 1**: Se X e Y são v.a independentes com fgm $M_X(t)$ e $M_Y(t)$. Seja $Z = X + Y$, então $M_Z(t) = M_X(t)M_Y(t)$.
- **Proposição 2**: Se X é uma v.a com fgm $M_X(t)$ e $Y = a + bX$, então $M_Y(t) = e^{at}M_X(bt)$.
- **Teorema da Continuidade de Levy**: Seja F_n uma sequência de funções de distribuições acumuladas com correspondentes funções geradoras de momentos $M_n(t)$. Seja F uma cdf com fgm $M(t)$. Se $M_n(t) \rightarrow M(t) \quad \forall t$ em um intervalo aberto contendo zero, então $F_n(y) \rightarrow F_Y(y) \quad \forall y$.
- Função geradora de momentos distribuição normal padrão $M_y(t) = e^{t^2/2}$.
- Exponencial $e^x = \lim_{n\rightarrow\infty} \left[1 + (x/n)\right]^n$.

Séries de Taylor

- Suponha que uma função $f(x)$ é derivável $(n+1)$ vezes em um intervalo contendo $x = x_0$.
- Expansão em **Série de Taylor** de $f(x)$ em torno de $x = x_0$ consiste em reescrever $f(x)$ da seguinte forma:

$$f(x) = f(x_0) + (x - x_0)\frac{df(x)}{dx}|_{x=x_0} + \frac{(x-x_0)^2}{2!}\frac{d^2 f(x)}{dx^2}|_{x=x0} + \frac{(x-x_0)^3}{3!}\frac{d^3 f(x)}{dx^3}|_{x=x_0} + \ldots + \frac{(x-x_0)^n}{n!}\frac{d^n f(x)}{dx^n}|_{x=x_0} + e_n(x)$$

onde o termo $e_n(x)$ é chamado de resíduo ou erro, e dado por

$$e_n(x) = \frac{(x-x_0)^{n+1}}{(n+1)!}\frac{d^{n+1} f(x)}{dx^{n+1}}|_{x=\epsilon}$$

sendo ϵ um valor entre x e x_0.

Teorema Central do Limite

- **Teorema Lindeberg-Levy**: Seja $Y_1, \ldots, Y_n$ v.a's iid com $\mathrm{E}(Y_i) = \mu$ e $\mathrm{V}(Y_i) = \sigma^2 < \infty$. Então,

$$\sqrt{n}\left(\frac{\bar{Y} - \mu}{\sigma}\right) \xrightarrow{D} Z \sim N(0,1), \quad \text{para} \quad n \rightarrow \infty.$$

- Isso significa que, para todo $y \in \Re$,

$$P(Y_n \leq y) \rightarrow \Phi(y) \quad \text{quando} \quad n \rightarrow \infty,$$

onde

$$\Phi(y) = \int_{-\infty}^{y} \phi(z) dz \quad \text{e} \quad \phi(z) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}z^2\right).$$

- Forma alternativa: $\overline{Y} \sim N(\mu, \sigma^2/n)$.

Teorema Central do Limite (versão multivariada)

- **Teorema Lindeberg-Levy multivariado**: Sejam $\boldsymbol{Y}_1, \ldots, \boldsymbol{Y}_n$ vetores aleatórios $p-$dimensionais com $\mathrm{E}(\boldsymbol{Y}_i) = \boldsymbol{\mu}$ e $\mathrm{V}(\boldsymbol{Y}_i) = [(\boldsymbol{Y}_i - \boldsymbol{\mu})(\boldsymbol{Y}_i - \boldsymbol{\mu})^\top] = \Sigma$, onde Σ é não-singular. Denote $\Sigma^{-1} = \Sigma^{-\frac{1}{2}}\Sigma^{-\frac{1}{2}}$. Então,

$$\sqrt{n}\Sigma^{-\frac{1}{2}}(\bar{\boldsymbol{Y}} - \boldsymbol{\mu}) \xrightarrow{D} \boldsymbol{Z} \sim N(\boldsymbol{0}, \boldsymbol{I}),$$

  onde $N(\boldsymbol{0}, \boldsymbol{I})$ denota uma distribuição normal multivariada com média zero e matriz de covariância identidade,

$$f(\boldsymbol{z}) = (2\pi)^{-p/2} \exp\left(-\frac{1}{2}\boldsymbol{z}^\top \boldsymbol{z}\right).$$

- Formas alternativas

$$\sqrt{n}(\bar{\boldsymbol{Y}} - \boldsymbol{\mu}) \sim N(\boldsymbol{0}, \Sigma)$$

$$\bar{\boldsymbol{Y}} \sim N(\boldsymbol{\mu}, n^{-1}\Sigma).$$

Teorema Central do Limite

- **Teorema Lindeberg-Feller:** Seja $Y_1, \ldots, Y_n$ v.a independentes $\mathrm{E}(Y_i) = \mu_i$ e $\mathrm{V}(Y_i) = \sigma_i^2 < \infty$. Defina, $\overline{\mu}_n = n^{-1}\sum_{i=1}^{n}\mu_i$ e $\overline{\sigma}_n^2 = n^{-1}\sum_{i=1}^{n}\sigma_i^2$. Suponha que

$$\lim_{n\to\infty} \max_i \frac{\sigma_i^2}{n\bar{\sigma}_n^2} = 0$$
$$\lim_{n\to\infty} \bar{\sigma}_n^2 = \bar{\sigma}^2 < \infty.$$

Então

$$\sqrt{n}\left(\frac{\bar{Y} - \bar{\mu}_n}{\bar{\sigma}_n^2}\right) \xrightarrow{D} Z \sim N(0,1).$$

- De forma equivalente,

$$\sqrt{n}(\overline{Y} - \bar{\mu}_n) \xrightarrow{D} N(0, \overline{\sigma}_n^2).$$

Teorema Central do Limite

- **Teorema Liapounov**: Sejam $Y_1, \ldots, Y_n$ v.a independentes $\mathrm{E}(Y_i) = \mu_i$ e $\mathrm{V}(Y_i) = \sigma_i^2 < \infty$. Suponha que
$$\mathrm{E}[|Y_i - \mu_i|^{2+\delta}] \leq M < \infty$$
para algum $\delta > 0$. Se $\bar{\sigma}_n^2$ é positiva e finita para todo n suficientemente grande, então
$$\sqrt{n}\left(\frac{\overline{Y} - \overline{\mu}_n}{\overline{\sigma}_n}\right) \xrightarrow{D} Z \sim N(0, 1).$$
- Equivalentemente,
$$\sqrt{n}(\overline{Y} - \overline{\mu}_n) \xrightarrow{D} N(0, \overline{\sigma}^2).$$
- Liapounov é equivalente a Lindeberg-Feller só que mais simples de entender e verificar na prática.

Resultados para manipular TCL's

- **Teorema Slutsky**: Sejam Y_n e Z_n sequência de v.a tais que $Y_n \xrightarrow{D} Y$ e $Z_n \xrightarrow{P} c$, onde Y é uma v.a e c é uma constante. Então, os seguintes resultados valem quando $n \rightarrow \infty$:
  - $Z_n Y_n \xrightarrow{D} cY$.
  - $\frac{Y_n}{Z_n} \xrightarrow{D} \frac{Y}{c}$ desde que $c \neq 0$.
  - $Y_n + Z_n \xrightarrow{D} Y + c$.

Exercício

1. Seja $Y_1, \ldots, Y_n$ v.a iid com $\mathrm{E}(Y_i) = \mu$ e $\mathrm{V}(Y_i) = \sigma^2 < \infty$.

   a) Mostre que $\overline{\sigma}^2 \xrightarrow{P} \sigma^2$ onde $\overline{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \mu)^2$.

   b) Obtenha a distribuição aproximada de $\overline{\sigma}^2$.

   c) Faça uma ilustração computacional da distribuição aproximada conforme o tamanho da amostra cresce. Use $n = 50, 250$ e 1000.

   d) Compare computacionalmente a distribuição aproximada com a distribuição exata apresentada na semana 3.